

Gli sviluppi effettuati sul software Eprints al CILEA

Zeno Tajoli

CILEA, Segrate

Abstract

L'articolo descrive gli sviluppi tecnici fatti per gli archivi gestiti con Eprints presso il CILEA

This article describes the technical improvements implemented in the two Eprint repositories managed by CILEA.

Keywords: Eprints, repository, biblioteconomia, medicina sociale.

Dati generali

Il CILEA gestisce due *repository* realizzati con il software Eprints [1], E-LIS e e-ms. E-LIS [2] è il maggior *repository* a livello internazionale che si occupi di biblioteconomia e di scienze dell'informazione, contiene più di 4500 testi ed è attivo dal gennaio 2003. Il *repository* e-ms [3] è nato nell'agosto 2005, per iniziativa dell'Istituto Italiano di Medicina Sociale, e accoglie sia le pubblicazioni dell'istituto sia materiali di esterni, sempre nell'ambito della medicina sociale e delle discipline correlate. Contiene circa 480 testi.

Per ambedue le installazioni è stata compiuta una serie di interventi di personalizzazione. Alcune modifiche del codice sono state applicate a entrambi i *repository*, altre sono specifici, come lo sono tutti gli interventi sull'interfaccia utente.

Sono stati anche sviluppati alcuni nuovi strumenti a linea di comando per aiutare nell'amministrazione del sistema.

Interventi sul codice condivisi

Per migliorare il funzionamento del software sono stati fatti degli interventi sul codice perl alla base del sistema.

Si è intervenuti sul modo con cui vengono creati i file di riepilogo per le liste di scorrimento. Nel software di base vengono creati nomi molto lunghi, perché si evita di usare caratteri non ASCII. Poiché il *file system*

del nostro server permettere di creare file i cui nomi siano in UTF-8, sono state tolte le routine che trasformavano nomi con caratteri UTF-8 in nomi con caratteri ASCII. Questo ha permesso di avere liste di scorrimento con nomi molto lunghi come la lista per nome di congressi.

Sono poi stati inseriti gli abstract nelle ricerche periodiche impostate dagli utenti, gli "alerts".

Un ulteriore intervento ha riguardato la funzionalità di ricerca che parte dalle citazioni bibliografiche. Si è fatto in modo che, se la citazione contiene un link web, il sistema porti direttamente l'utente a quel link. Se invece la citazione non contiene link, si cerca di estrarre i singoli dati presenti e di impostare una ricerca all'interno dell'archivio stesso. Se non si trova niente, si lancia una ricerca sul servizio Paracite [4]

Piccoli interventi sono stati fatti per gestire al meglio i caratteri non Latin-1. Questa problematica è evidente soprattutto in E-LIS, che riceve testi e metadati in molti alfabeti (Latino base, Latino Esteso, Greco, Cinese, ecc.), ma i cambiamenti sono operativi per entrambi gli archivi.

Interventi sul codice non condivisi

La maggior parte degli interventi non condivisi riguarda E-LIS. Per prima cosa, si è proceduto alla creazione di una pagina di riepilogo di tutti i metadati, accessibile via web

nella forma “./fullmetadata.html”. Questa pagina viene usata dai redattori per controllare la qualità dei metadati, senza doversi far riconoscere dal sistema.

Un secondo intervento è stato fatto sull'interfaccia riassuntiva che fa vedere i documenti che devono essere controllati dai redattori prima di essere resi visibili al pubblico di Internet. Si è aggiunta una colonna che indica la lingua del full-text collegato ai metadati.

Sempre per i redattori, si è aggiunto tra i metadati un campo *country*, da compilare per poi dividere i vari lavori tra i redattori dei diversi paesi.

Sempre per i redattori, si è pensato di salvare una copia della mail inviata a un autore quando respingono un lavoro per le più diverse motivazioni. Agli autori è stata invece data la possibilità di ricevere un mail quando il proprio lavoro diventa disponibile sul web. Queste due funzionalità sono state provate e testate e saranno a breve operative.

A e-ms invece è stata aggiunta la possibilità di derivare i metadati di una pubblicazione direttamente da PubMed.

Interventi sull'interfaccia utente

Per quanto riguarda e-ms, si è lavorato molto sull'interfaccia utente con XHTML e CSS. In particolare, la versione italiana è ora conforme in maniera completa agli standard del W3C e si è migliorata molto l'accessibilità. Inoltre, è stata fatta ad hoc una nuova traduzione, considerando che il sito non viene usato solo da personale accademico, ma anche da pazienti e da medici.

Per E-LIS, nella homepage è stato inserito un modulo, realizzato da terzi e opportunamente modificato, per mostrare le ultime novità.

Inoltre, sono state aggiunte le statistiche d'uso dei singoli testi e alcune a livello generale. Il sistema usato si rifà a quanto fatto dall'Università della Tasmania [5], ma è stato quasi completamente riscritto in perl, per rendere le pagine da dinamiche a statiche, poiché il server non riusciva a reggere la generazione dinamica. Si ritiene che questo sia dovuto all'utilizzo di MySQL 3.23, mentre l'applicativo per le statistiche è stato sviluppato con MySQL 4.x, che ha una funzionalità di *caching* non presente nella versione 3.x.

Nuove utilità di sistema

Gli *script* che sono stati realizzati servono per fare operazioni non disponibili via web, come:

- controllare se i soggetti inseriti sono conformi alla classificazione impostata nel sistema;
- controllare se sono presenti i metadati considerati obbligatori;
- eliminare gli utenti che non hanno confermato la loro registrazione;
- eliminare gli utenti che non hanno mai compiuto operazioni a un anno dall'iscrizione.
- elencare gli e-mail degli utenti;
- eliminare i caricamenti per identificativo interno (eprintid).

Bibliografia

- [1] URL: <http://www.eprints.org/software/>
- [2] URL: <http://eprints.rclis.org/>
- [3] URL: <http://e-ms.cilea.it/>
- [4] URL: <http://paracite.eprints.org/>
- [5] URL: <http://files.eprints.org/25/>